

**SYSTEM PERFORMANCE ACCOUNTABILITY SYSTEM:
REVISED ACCOUNTABILITY INDEX**

BACKGROUND

The State Board approved the provisional State Board of Education (SBE) Accountability Index and proposed recognition system in May 2009. The provisional SBE Accountability Index provides ways to examine all schools and districts based on a variety of different outcomes and indicators. Schools and districts can use the index visually to determine where they have made significant improvements as well as where achievement gaps exist for their students based on reading, writing, math, science and for high schools on extended graduation rates.

The Office of Superintendent of Public Instruction (OSPI) and the SBE plan to recognize schools for high achievement and improvement in May 2010, using the SBE Provisional Accountability Index with some OSPI additions. However, the new SBE Accountability Index will not be used for identifying low achieving schools unless the SBE and OSPI are successful in obtaining a waiver from the U.S. Department of Education or making changes in the No Child Left Behind (NCLB) Reauthorization. The SBE staff will work with OSPI in 2010 to request that the federal government consider our SBE Accountability Index.

For now, we are following the new federal school improvement rules to identify lowest achieving schools. These rules are better than what exists under NCLB to identify a school in improvement. However, the federal school improvement rules do not allow all districts and schools to look at their progress, or lack of progress over time, unless they are identified as part of the lowest five percent of persistently low achieving schools. Nor do the federal school improvement rules include subjects beyond math and reading.

The ultimate goal of the final SBE Accountability Index will be to replace the federal method of determining Annual Yearly Progress (AYP) as well as recognition and continuous improvement for all schools if the federal government permits Washington to use the Accountability Index. Staff sought feedback from stakeholders on additional improvements needed to the provisional SBE Accountability Index in the summer and fall of 2009. Most of the feedback centered on creating an easy to understand paper on the index and adding a way to address different groups of students' achievement.

Staff worked with writer, Jill Severn, and BERK & Associates graphics to develop a layperson's guide to the provisional SBE Accountability Index. See Attachment A "Creating A Better Accountability Quality Rating System" paper and Attachment B "Accountability Index: An Introduction" graphic.

SBE consultant, Pete Bylsma, has revised the provisional SBE Accountability Index for final Board review and action at the January 13-14, 2010 meeting. This item was held over from the November 2009 Board meeting.

POLICY CONSIDERATION

The SBE consultant, Pete Bylsma, has revised the SBE Provisional Accountability Index approved in May 2009, to include performance subgroups of students as part of the new Accountability Index. See the tab titled "Index Updates" for the "Executive Summary: Washington's new Accountability Index and "Recommendations for Accountability Student Groups".

EXPECTED ACTION

Staff recommends that the Board review and approve the revised SBE Accountability Index that incorporates an analysis of student subgroups through an additional matrix, to determine AYP based on the "Executive Summary: Washington's new Accountability Index" and "Recommendations for Accountability Student Groups." SBE staff will then work with OSPI to seek federal approval for the use of this final SBE Accountability Index.

Staff also recommends the Board approve a specific recommendation to the Office of Superintendent of Public Instruction to post the results of the Washington Language Proficiency Test results for English Language Learners, by district, on the OSPI Web site for the 2009-2010 school year.

Creating A Better Accountability Index Rating System

No federal law has a nobler purpose than the 2001 No Child Left Behind Act (NCLB). Its goal is to ensure that all American children get the quality education they need to thrive and to sustain the democracy and prosperity of our country.

More specifically, NCLB was designed to hold public schools accountable for doing a better job of educating children of color, children from low-income families, children who are English language learners, and children with disabilities. These are the groups of children who have most often been left behind in the past. For far too long, these children have been trapped in a system and in schools that are under-resourced and unprepared to meet their educational needs.

To hold schools accountable for improving the education of all children, NCLB created a system of quality ratings for schools and sanctions for schools that do not meet specific goals for student achievement. This rating system sets a very ambitious goal: by 2014, all U.S. students would pass state tests in reading and math.

NCLB requires that schools make Adequate Yearly Progress (AYP) towards this goal by scoring at progressively higher levels in a complex matrix of categories that includes reading and math scores for nine groups of students (five ethnic groups, low-income students, English language learners, and students in special education). The matrix also includes additional factors, such as high school graduation rates and the number of unexcused absences. For each of the student groups, scores are generated whenever there are 30 or more students in that subgroup in a school.

When schools that receive federal Title I funds do not meet the AYP goals established by NCLB, they enter a “step of improvement.” (Title I funds are directed to schools with a large proportion of low-income students.) This means that schools are required to take certain actions, such as giving all students the opportunity to transfer to other schools or providing extra tutoring at no cost to students’ families. If schools still aren’t able to make improvements and raise their scores, more drastic sanctions are required, including state intervention in persistently low-achieving schools.

This system of quality rating and progressive sanctions was intended to focus schools on meeting the educational needs of all students. But critical flaws in the complex system of ratings and sanctions produce misleading results and unintended consequences. Specifically, the flaws of the AYP system include:

- Where there are fewer than 30 students in a measured sub-group, that group’s educational progress is not counted. This often means small schools are excluded, as well as small or isolated groups of students of color, English language learners, and students in special education.
- Sanctions apply only to schools receiving federal Title I funds.
- Sanctions are not tailored to address the specific needs of local schools and districts and do not result in needed improvements.
- The system is so complex that educators, parents, and the public are often confused rather than informed.
- The current federal system only measures student achievement in reading and math.
- The current system does not recognize or reward improvements in student achievement.
- Washington has higher academic standards than many other states, so the bar is higher for our schools. As a result, the current federal system classifies too many of our schools as “in need of improvement” and subject to sanctions.

Creating a new and better quality rating system for schools

There is growing momentum all across the country for substantial change to the accountability provisions of NCLB. Pending Congressional reauthorization of the NCLB Act presents an opportunity for change. At the same time, the Obama administration has signaled its understanding of the need for change by proposing regulations requiring a vastly simpler set of metrics that schools must meet to win federal "Race to the Top" grants for school improvement efforts.

The Washington State Legislature directed the State Board of Education (SBE) to develop a simpler, fairer, more informative accountability index to identify schools that deserve recognition and those that need additional support or intervention to improve.

This new Accountability Index could, if approved by Congress, replace the current NCLB requirements. The federal Department of Education could also grant a waiver that allows Washington State to use this system in place of the one required by NCLB.

The SBE and the Office of Superintendent of Public Instruction (OSPI) intend to use this new, locally created Accountability Index to assess the quality of our state's schools beginning in 2010. Even though this system cannot yet be used in place of federal requirements, it is needed to give us a more accurate picture of how our schools are progressing and where improvement is needed.

Washington's Accountability Index

Principles that guided the development of the Accountability Index:

- Use existing data.
- Rely on multiple measures.
- Include assessment results from all grades and subjects tested statewide.
- Be fair, reasonable, and consistent.
- Be valid and accurate.
- Use consistent rules for both schools and school districts.
- Apply to as many schools and districts as possible.
- Use familiar concepts when possible.
- Provide multiple ways to reward success.
- Be flexible enough to accommodate future changes.

The Accountability Index is based on a set of *five outcomes* and *four indicators*. The outcomes are the *results of state tests in reading, writing, math, and science*.

These five outcomes are examined using four indicators:

- Achievement of non-low income students.
- Achievement of low-income students.
- Achievement of all students compared to those in schools or districts with similar demographic characteristics.
- Improvement in student achievement over the previous year.

Table 1: Matrix of Accountability Measures

INDICATORS	OUTCOMES				
	Reading	Writing	Math	Science	Ext. Grad. Rate
Achievement of non-low income students.					
Achievement of low income students.					
Achievement vs. peers.					
Improvement from the previous year.					

How the rating system works

Ratings used in the index range from 1 to 7 and are based on a school’s performance on each of the outcomes and indicators.

Achievement is measured by the percentage of students who meet the state’s academic standards in reading, writing, math, and science. The highest rating (7) applies to schools or districts where at least 90 percent of students meet academic standards; the lowest rating (1) is applied to schools where fewer than 40 percent of students do so. Separate ratings are given for non-low income and low income students. Ratings are also given for the extended graduation rates of these two student groups.

A *Learning Index* is used to calculate Achievement vs. Peers and Improvement.

The *Achievement vs. Peers* measure is determined by predicting the average level of achievement that would occur in schools and districts with similar student characteristics – that is, similar percentages of students who are from low income families, are English language learners, are highly mobile, or who are in special education or gifted programs. Ratings are based on how far a school or district is above or below the predicted level.

The Learning Index measures Achievement vs. Peers in reading, writing, math, and science. This index is based on the percentage of students who score at each proficiency level on Washington’s standards-based tests:

- 4 = Exceeds Standards
- 3 = Meets Standards
- 2 = Partially Meets Standards
- 1 = Well Below Standards

The Learning Index averages all the student results. The score is similar to a grade point average. Thus, if a school’s Learning Index is *above what is predicted* by .20 (similar to a difference in grade point average of 2.50 and 2.70) the school receives a rating of 7. This score recognizes that the school outperformed those with similar student characteristics. Scores in the middle of the range (-.05 to .05) show that a school or district is performing on par with peers with similar student characteristics. Scores at the bottom of the range indicate that a school or district has fallen behind its peers.

A distinctive feature of the Learning Index is that it includes a measure for students who *exceed standards*. In the current federal AYP system, there is no recognition or reward for having students meet the highest level of achievement.

The Learning Index is also used to measure “*Improvement*.” The score is the amount of change that took place in the Learning Index from the previous year. Higher ratings are given when the Learning Index increases and lower ratings are given when they decline.

RATING SYSTEM

	READING	WRITING	MATH	SCIENCE	EXT. GRAD. RATE	
ACHIEVEMENT FOR NON-LOW INCOME STUDENTS	<u>% MET STANDARD</u> <u>RATING</u>				<u>RATE</u>	<u>RATING</u>
	90 - 100%7				> 95%.....7	
	80 - 89.9%6				90 - 95%6	
ACHIEVEMENT FOR LOW INCOME STUDENTS	70 - 79.9%5				85 - 89.9%5	
	60 - 69.9%4				80 - 84.9%4	
	50 - 59.9%3				75 - 79.9%3	
	40 - 49.9%2				70 - 74.9%2	
	< 40%1				< 70%1	
ACHIEVEMENT VS. PEER SCHOOLS¹	DIFFERENCE IN				DIFFERENCE	
	<u>LEARNING INDEX</u> <u>RATING</u>				<u>IN RATE</u>	<u>RATING</u>
	> .207				> 127	
	.151 to .206				6.1 to 126	
	.051 to .155				3.1 to 65	
	-.05 to .054				-3 to 34	
	-.051 to -.153				-3.1 to -63	
	-.151 to -.202				-6.1 to -122	
< -.201				< -121		
IMPROVEMENT	CHANGE IN				CHANGE	
	<u>LEARNING INDEX</u> <u>RATING</u>				<u>IN RATE</u>	<u>RATING</u>
	> .157				> 67	
	.101 to .156				4.1 to 66	
	.051 to .105				2.1 to 45	
	-.05 to .054				-2 to 24	
	-.051 to -.103				-2.1 to -43	
	-.101 to -.152				-4.1 to -62	
< -.151				< -61		

¹ This indicator controls five student variables beyond a school’s control: the percentage of low-income, ELL, special education, gifted, and mobile students. Scores are the difference between the actual level and the predicted level of the Learning Index, which accounts for the performance of students who have not yet reached state standard and those who exceed state standards. Scores above 0 are “beating the odds” and negative scores reflect performance below the predicted level.

High schools and school districts have 20 measures, while elementary and middle or junior high schools have only 16, because they do not have graduates.

School districts are assessed using the same indicators, outcomes, and criteria as schools. The results are based on district-wide data for all grades. The financial data are used in the peers’ analysis to control the amount of total operating expenditures per student.

Measuring progress in closing achievement gaps

For subgroups of students that are of special concern, a simplified matrix measures achievement in reading, math, and the extended graduation rate (which includes students who take more than four years to graduate). This matrix uses the same concepts as this Accountability Index to measure achievement and improvement among English language learners, students in special education, African-American, Latino, Native American, Asian, and low-income students. The Washington State Accountability Index adds two more subgroups to make the system more comprehensive: Pacific Islanders and multi-racial students.

HYPOTHETICAL EXAMPLE OF A SCHOOL'S SUBGROUP RESULTS

Ratings in **red** represent declines in student achievement from the previous year; ratings in **green** represent improvements over the previous year.

Subgroup	READING			MATH			EXT. GRAD. RATE			Average rating	Change from previous year
	Met Std. (All stud.)	Peers	Improve.	Met Std. (All stud.)	Peers	Improve.	Met Std. (All stud.)	Peers	Improve.		
American Indian	4	4	4	1	5	4	1	4	4	3.44	0.33
Asian	6	4	5	4	5	4	5	4	6	4.78	0.56
Pacific Islander	4	4	4	1	5	4	1	4	4	3.44	0.22
Black	3	3	3	1	3	5	1	3	2	2.67	-1.00
Hispanic	3	4	4	1	4	4	1	4	4	3.22	-0.11
White	5	4	4	3	3	4	3	4	4	3.78	-0.22
Multi-racial	4	4	4	2	4	4	2	4	4	3.56	-0.22
Special education	2	4	4	1	4	4	1	2	1	2.56	-1.22
ELL	1	5	5	1	4	4	1	3	3	3.00	-0.11
Low income	4	5	5	1	5	4	1	4	4	3.67	0.22
Average	3.6	4.1	4.2	1.6	4.2	4.1	1.7	3.6	3.6	3.41	-0.16

The Accountability Index creates tiers of schools' performance

In the new matrix, schools are sorted into five tiers. The following table, applied to schools using data from 2007, provides essential information about student achievement and school performance. It shows educators, parents and the public where the school is strong, and where improvement is needed. Tables like this could be publicly available on OSPI's website.

Table 4: School Assessment via the Accountability Index

INDICATORS	OUTCOMES					AVERAGE
	READING	WRITING	MATH	SCIENCE	EXTENDED GRADUATION RATE	
Achievement of non-low income students	5	6	3	1	5	4.00
Achievement of low income students	4	4	1	1	7	3.40
Achievement vs Peers	2	2	2	2	6	2.80
Improvement from the previous year	1	4	1	4	7	3.40
AVERAGE	3.00	4.00	1.75	2.00	6.25	3.40

Special features of the Accountability Index

The Accountability Index builds in flexibility to account for special cases and circumstances:

- Alternative schools, which often help students with the greatest need and therefore often have lower overall student scores, will receive additional scrutiny to ensure they are using best practices.
- Results for English language learners can be excluded for three years (rather than the federal requirement of only one year) to give these students a realistic amount of time to acquire the language skills necessary for tests that require reading and writing in English.
- To provide more information about how quickly students are learning English, OSPI will include test results from the Washington Language Proficiency Test on its School Report Card website.
- Schools that perform at very high levels, and therefore don't have room for significant improvement on the Index, will not be penalized because the "improvement" rating won't be counted for them.

How the Accountability Index can identify and rank schools and districts that deserve recognition for excellence and those that are not making adequate yearly progress (AYP)

If the Accountability Index is approved by a waiver from the U.S. Department of Education or incorporated into the NCLB when it is reauthorized by Congress, it could guide both the classification and the remedies for low-performing schools and districts.

Even without such approval, OSPI and SBE intend to use the Accountability Index to generate more accurate information about how Washington's schools are performing and to identify schools that deserve special recognition for excellence.

Each fall, OSPI will compute the Accountability Index and sort all schools and districts into five tiers. Awards and recognition will be bestowed on our best schools and districts and on those who "beat the odds" by narrowing the achievement gap.

If federal approval is obtained, the Accountability Index will guide what all persistently low-performing schools (not just those that receive federal Title I funds) must do to improve.

The first time a school falls into the bottom tier of the Index, it will be considered a "warning" year. Schools and districts that do not make AYP two years in a row will undergo an in-depth review to more closely identify trends, resource and staffing issues, and other data that will determine whether they should move into "improvement" status and be required to take certain actions.

When a school or district has not made AYP for four years, the state may determine that it would benefit from a significant amount of additional support and classify it as in need of "Voluntary Action." This status will offer – but not require – the school or district to participate in OSPI's School and/or District Improvement Program, which provides outside consultants, professional development, coaching, and other evidence-based assistance to accelerate improvement.

If improvement does not occur after two more years (regardless of whether the school or district participates in offered state programs), it will move to "Required Action" status, and a binding corrective action plan will be established between the school or district and the SBE. (This step will require legislative approval, since current state law does not allow required state intervention in local schools or districts.)

Creating a system of excellent and equitable education for all students

Washington's Accountability Index is one part of an emerging state system that can end the persistent low performance of some schools and ensure a focus on improved performance for all schools and districts. It will provide accurate, fair, easy-to-understand information on school and district performance and progress. That information can guide local efforts to improve and target customized state assistance and intervention when local efforts don't produce the needed results.

Most important, the Accountability Index, combined with a new system for state intervention and OSPI's well-regarded School and District Improvement Programs, can change the future of students who are languishing in schools that are not providing them with the education they urgently need.

Attachment B

Insert PDF when printed for “Accountability Index: An Introduction” graphic.

Executive Summary

Washington's New Accountability Index Final Report to the State Board of Education

January 2010

Pete Bylsma, EdD, MPA
BYLSMAPJ@COMCAST.NET

*The full report is available on the
State Board of Education Web site at
<http://www.sbe.wa.gov/spa.htm>.*

CREATING THE ACCOUNTABILITY INDEX

The Legislature requires the State Board of Education (SBE) to develop a statewide accountability system to help improve academic performance among all students. SBE was required to “adopt objective, systematic criteria” to identify schools and districts for recognition and for receiving additional state support. The 2009 Legislature required the Board to develop an index for such purposes. To meet this requirement, the Board has developed a provisional Accountability Index to sort schools and districts into different “tiers” based on multiple measures. The Board believes the index plays a key role in providing feedback about the status of education reform in schools and districts and in supporting continuous improvement efforts. Schools and districts in most need will be eligible to receive more significant state support and will be required to participate in a state system of support if initial offers of more support are not accepted and substantial improvement does not occur after several years. The creation of the index comes at a time when changes in the state’s assessment and data systems and at the U.S. Education Department provide an opportunity to consider new accountability ideas. However, the recommendations made under this index cannot be used by the state to identify struggling schools for Adequate Yearly Progress (AYP) until the U.S. Education Department approves it through either a waiver or through the reauthorization of No Child Left Behind (NCLB) Act that allows this system.

Various principles guided the development of the index. The index needs to (1) be transparent and simple to understand, (2) use existing data, (3) rely on multiple measures, (4) include assessment results from all grades and subjects tested statewide, (5) use concepts of NCLB and its AYP system when appropriate, (6) be fair, reasonable, and consistent, (7) be valid and accurate, (8) apply to both schools and districts, (9) apply to as many schools and districts as possible, (10) use familiar concepts when possible, (11) rely mainly on criterion-referenced measures instead of norm-referenced measures, (12) provide multiple ways to reward success, and (13) be flexible enough to accommodate future changes.

The provisional index is based on how schools and districts perform on a set of five outcomes and four indicators. The five outcomes are the results of state assessments in four subjects (reading, writing, mathematics, science) and the “extended” graduation rate (for high schools and districts). These five outcomes are examined using four indicators: achievement of (1) non-low income students, (2) low-income students, (3) all students compared to those in similar schools/districts (controlling for the percentage of students who are learning English, have a disability, live in a low-income home, are mobile, and are designated as gifted), and (4) the level of improvement from the previous year. The results of the 20 measures form the 5x4 matrix shown in Table 1.

Table 1: Matrix of Accountability Measures

INDICATORS	OUTCOMES				
	Reading	Writing	Math	Science	Ext. Grad. Rate
Achievement of non-low income students					
Achievement of low income students					
Achievement vs. peers					
Improvement from the previous year					

Each cell of the matrix is rated on a 7-point scale (from 1 to 7) using fixed benchmarks. Each of the four subjects is rated using the same set of benchmarks across the entire school (i.e., all subjects have the same set of benchmarks, and the assessment results are the aggregate totals for students in all the tested grades). **The index is the simple average of all the ratings** and ranges from 1.0 to 7.0. High schools and districts have 20 measures, while elementary and middle/junior high schools have only 16 measures because they do not have graduates. Table 2 shows how each of the five outcomes are measured using the four indicators and the benchmarks that produce the ratings.

Table 2: Benchmarks and Ratings for Outcomes and Indicators

	READING	WRITING	MATH	SCIENCE	EXT. GRAD. RATE ¹
ACHIEVEMENT (NON-LOW INCOME)	<u>% MET STANDARD</u> <u>RATING</u> 90 - 100% 7 80 - 89.9% 6 70 - 79.9% 5 60 - 69.9% 4 50 - 59.9% 3 40 - 49.9% 2 < 40% 1				<u>RATE</u> <u>RATING</u> > 95..... 7 90 - 95% 6 85 - 89.9% 5 80 - 84.9% 4 75 - 79.9% 3 70 - 74.9% 2 < 70%..... 1
ACHIEVEMENT (LOW INCOME)					
ACHIEVEMENT VS. PEERS²	DIFFERENCE IN <u>LEARNING INDEX</u> <u>RATING</u> > .20..... 7 .151 to .20..... 6 .051 to .15 5 -.05 to .05 4 -.051 to -.15 3 -.151 to -.20 2 < -.20 1				DIFFERENCE <u>IN RATE</u> <u>RATING</u> > 12..... 7 6.1 to 12 6 3.1 to 6 5 -3 to 3..... 4 -3.1 to -6 3 -6.1 to -12 2 < -12..... 1
IMPROVEMENT³	CHANGE IN <u>LEARNING INDEX</u> <u>RATING</u> > .15..... 7 .101 to .15..... 6 .051 to .10..... 5 -.05 to .05 4 -.051 to -.10 3 -.101 to -.15 2 < -.15 1				CHANGE <u>IN RATE</u> <u>RATING</u> > 6..... 7 4.1 to 6 6 2.1 to 4 5 -2 to 2..... 4 -2.1 to -4 3 -4.1 to -6 2 < -6..... 1

Note: Assessment results are the combined results from both the state content assessments (e.g., WASL) and the WAAS (assessments for students with disabilities) from all grades.

¹ This outcome only applies to schools and districts that are authorized to graduate students.

² This indicator adjusts the outcomes using statistical methods (multiple regression) to control for five student characteristics beyond a school’s control: the percentage of low-income, ELL, special education, gifted, and mobile students. (Mobile students are those who are not continuously enrolled from October 1 through the entire testing period.) Scores are the difference between the actual level and the predicted level of the Learning Index. Scores above 0 are “beating the odds” and negative scores are below the predicted level. Separate analyses are conducted for schools for each of the four assessments for each type of school (elementary, middle, high, multiple grade levels). District calculations also control for the level of current expenditures per pupil (adjusted for student need).

³ Measured in terms of the change in the Learning Index from the previous year.

The proposed system holds *districts* accountable using the same indicators, outcomes, and criteria that are used for schools. The results are based on districtwide data for all grades rather than being

disaggregated by grade bands (elementary, middle, high). In addition, financial data are used in the “peers” analysis to control for the amount of total operating expenditures per pupil (adjusted for student need).

PRELIMINARY RESULTS

Schools and districts fall into five tiers based on the index score. In-depth analyses of the data and conditions occurs for schools and districts that do not make AYP two years in a row to see if they merit further support.

Table 3 shows the ranges for the tier assignments and the number of schools and districts that would have been placed in each tier in 2007 using the above criteria. Schools show a greater range than districts—far fewer districts were in the top and bottom tiers compared to the school results. The 228 schools in the Struggling tier enrolled 74,000 students (1 in 14 students statewide). Of the schools in this tier, 98 (43%) were alternative schools or served other special populations, and enrolled a total of 12,400 students. The 130 “regular” schools in the Struggling tier enrolled 61,600 students. Over the 2-year period, 149 schools (7.4%) had an average index below 2.50, and 89 were regular schools that enrolled approximately 39,000 students (roughly 4% of statewide enrollment). The 17 districts in the Struggling tier in 2007 tended to be rather small, averaging roughly 1,000 students. However, some larger districts had many schools in a struggling tier—17 districts had at least two regular schools and four districts had at least five regular schools with a 2-year index average below 2.50.

Table 3: Tier Ranges and Preliminary Results (2007)

Tier	Index Range	# of Schools	% of Schools	# of Students ¹	# of Districts	% of Districts	# of Students ¹
Exemplary	5.50 – 7.00	81	4.0%	28,650	1	.3%	360
Very Good	5.00 – 5.49	131	6.5%	64,500	9	3.1%	31,500
Good	4.00 – 4.99	591	29.4%	314,700	87	29.9%	278,500
Fair	2.50 – 3.99	980	48.7%	523,000	177	60.8%	692,500
Struggling	1.00 – 2.49	228	11.3%	74,000	17	5.8%	17,500

¹Approximate number (some schools did not provide enrollment data).

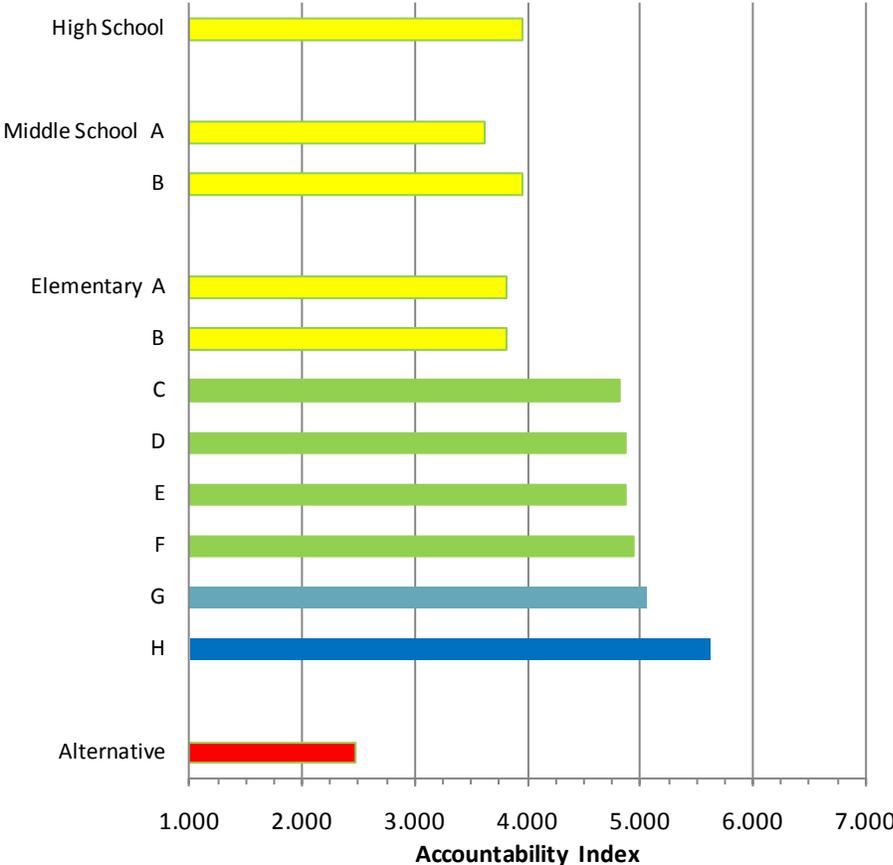
Table 4 provides an example of the ratings for an actual high school and how the average of the individual ratings generates the index/tier assignment. The school’s average rating of 3.40 is the index score, which puts the school in the middle of the Fair tier. The average ratings have been color-coded so the overall results can be seen at a glance. These types of results should be made public on the state Web site (the format for presenting the results must be determined). Results presented in this “dashboard” give policymakers, educators, and the public a quick snapshot of where a school is strong and weak, and it provides transparency about how the index is determined.

Table 4: “Actual” High School, 2007

Indicator	Reading	Writing	Math	Science	Grad Rate	Average
Non-low inc. ach.	5	6	3	1	5	4.00
Low-inc. ach.	4	4	1	1	7	3.40
Ach. vs. peers	2	2	2	2	6	2.80
Improvement	1	4	1	4	7	3.40
Average	3.00	4.00	1.75	2.00	6.25	3.40 ← INDEX

Various charts can illustrate district results as well. Figure 1 shows an example of how the index could be shown for each school in a district. In this example (an actual district), one school reached the Exemplary tier.

Figure 1: Accountability Results in “Actual” District, 2008



HANDLING SPECIAL CASES

The accountability system needs to be flexible enough to accommodate some special situations. These include holding alternative schools accountable using additional data, excluding some ELL results from the index calculations, and not counting the improvement cells when achievement is at very high levels.

Holding **alternative schools** accountable poses unique challenges. Many alternative schools exist in the state, and they vary greatly in their focus, structure, and clientele. Most are relatively small (total enrollment is less than 4% statewide), and more than half serve at-risk students in grades 9-12. Some believe these schools have taken on more challenging students while allowing traditional schools to generate better outcomes with their remaining students. On the other hand, some alternative schools offer special programs for students who are not at-risk and who meet rigorous academic requirements for admission. Some are considered “alternative” because they do not use a normal school approach, and a growing number of schools serve students through digital learning via the Internet. Parent Partnership Programs are a type of “school” where parents are the primary instructor and the district provides instructional support. Some target special student populations (e.g., special education, gifted, ELL). Given this variation, no “peer” indicator is computed for these schools.

Some alternative schools intentionally target student populations facing significant challenges and therefore are more likely to be in the Struggling tier. These schools should receive the normal index score based on calculations used for all schools. Alternative schools that do not make AYP in two consecutive years should be examined more closely to determine if they are using research-based best practices and showing progress. Areas for improvement should be identified and should be the focus of analysis if the alternative school does not make AYP again in the future.

Results for ELL students are currently included in AYP calculations in the student’s second year of enrollment in a U.S. public school. OSPI has asked the U.S. Education Department to exclude ELL results until a student has been enrolled in a U.S. public school for three years or until the student achieves an advanced level of English proficiency on the Washington Language Proficiency Test (WLPT), whichever comes first.¹ This request is based on research that shows it takes many years for an ELL student to acquire “academic” proficiency in English and because the student must be able to read and write English to understand and respond to each test item. Moreover, testing students who do not understand English violates widely-adopted testing standards because of threats to validity and mistreatment of human subjects. However, the Department has denied OSPI’s repeated request to use this policy.

Nevertheless, to improve the validity of the accountability system, computing the index should exclude the results for ELL students who are in their first three years of enrolling in a U.S. public school for any test that requires reading and writing in only English.² In addition, OSPI should begin reporting WLPT results on its Report Card in a way that allows educators, parents, and other stakeholders to monitor the progress ELLs make in terms of learning English and meeting state content standards. Finally, all ELLs should be required to take the state assessments after their first year of enrollment, and OSPI should analyze the various content assessments and WLPT results to determine the extent to which ELLs are on track to meet state standards. (These results may be used to determine AYP, as discussed below.)

Most ELL results would still be included in the accountability index, even with this “extended exclusion,” because (1) most ELLs enter school in kindergarten and have attended school for three years before taking state assessments for the first time in grade 3, and (2) most ELLs enrolled in the assessed grades (3-8 and 10) reach the advanced level of the WLPT. As a result, the exclusion has little impact on the index results. Nevertheless, the combination of recommendations improves the validity of the accountability system and provides more information about the progress of ELLs.

Schools and districts that perform at very high levels are not able to improve much from the previous year. To avoid “penalizing” these schools for a lack of improvement, the ratings for this indicator should not be included in the index calculations under certain conditions. Without this

¹ The composite score from the annual Washington Language Proficiency Test (WLPT) reflects proficiency in reading, writing, speaking, and listening. The three-year exemption period reflects the views of most stakeholders and is the average time required for ELL students to exit the program.

² The math and science tests were available in Spanish and Russian for the first time in 2009 but responses must still be made in English. There is no requirement to include the results of ELLs in their first three years if they reach Level 3 (advanced English) on the WLPT because data show these students do not yet know enough English to meet standard on the content assessments. When data become available on ELLs’ initial English level, the exclusion period should vary based on their incoming English ability. The recommended exclusion period is 2 years for students with advanced English (Level 3), 3 years for students with intermediate English (Level 2), and 4 years for students with limited/beginning English (Level 1).

policy, schools/districts with nearly all of their students achieving Level 4 on an assessment and graduating nearly all their students would not be able to achieve a rating above 4 (little or no improvement). Specifically, the improvement indicator should be excluded when computing the index whenever a Learning Index reaches 3.85 out of 4.00 and remains at or above that level for two consecutive years. (A school or district needs to improve by more than .15 to receive a rating of 7, which is impossible when their Learning Index reaches 3.85.) The first year the Learning Index falls above 3.85, a school/district would get a rating based on its improvement. If the Index stays at or above 3.85, the maximum rating is not possible and the indicator should not be calculated.³ The same policy applies to the extended graduation rate outcome (when the rate reaches or exceeds 94% in two consecutive years.⁴

INTEGRATING THE FEDERAL AND STATE ACCOUNTABILITY SYSTEMS

No Child Left Behind requires accountability for nine groups of students in reading, math, and one more indicator.⁵ Accountability for performance at the student group level is widely viewed as a positive feature of the law. Federal law also requires states to have a single accountability system. However, stakeholders across Washington believe the current federal system is overly complex and that the AYP results do not provide an accurate picture of school and district quality. As a result, stakeholders have provided advice on how elements of the provisional Accountability Index could be used to make AYP decisions. They also suggested changes in the consequences schools and districts face when they do not make AYP and when they make AYP while in improvement status.

Determining AYP

The following rules are recommended to hold schools and districts accountable for performance of various student groups.

- Hold the **All students** group accountable using the Accountability Index when there are at least 4 rated cells in the matrix.⁶ Specifically, *schools and districts with a 2-year average Accountability Index below 3.00 AND an index that declines two years in a row do not make AYP*. Using the Index in this way for AYP provides consistency in the accountability measure, and the required level is easy to understand and identifies a reasonable number of schools.
- Hold **subgroups** accountable using a separate modified matrix that uses the same concepts as the Accountability Index. Two more subgroups (Pacific Islanders, multi-racial) should be added to provide more complete coverage. However, only the outcomes used for federal accountability—reading, math, and the extended graduation rate—should be used, and the two income-related indicators should be combined. A “row average” should then be calculated for each subgroup. *Schools and districts do not make AYP if any row average declines two years in a row.*

³ Of the schools and districts with reportable data (at least 10 students assessed), two schools reached this level on the Learning Index in 2006 (one in both reading and writing and the other in writing), but no district reached this level in 2006 and no school or district reached this level in any subject in 2007.

⁴ Of the schools with graduation data, 11% had a rate that was at least 94% in two consecutive years.

⁵ The nine groups are “all” students, five race/ethnic groups, two program groups (students with disabilities and English language learners), and students from low income families. In Washington, the unexcused absence rate is the additional indicator at the elementary and middle school levels, and the extended graduation rate is the additional indicator for high schools.

⁶ Schools with fewer than 4 rated cells should submit an improvement plan to OSPI for review.

Table 5 gives an example for a hypothetical high school with at least 10 students in each subgroup (very few schools have at least 10 students in every group). Ratings are based on the performance of each group in three outcomes (reading, math, extended graduation) and three indicators (achievement of all students, achievement vs. peers, and improvement).⁷ In this example, six groups had a row average in 2009 that was less than the row average in 2008. If the row average in 2010 declines again for any of these groups, the school would not make AYP in that group. Colors are used to highlight ratings that are better or worse than the previous year.

Table 5: 2009 Results, Hypothetical High School

2009 Subgroup	READING			MATH			EXT. GRAD. RATE			Average rating	Change from previous year
	(all st.)	Peers	Improve.	(all st.)	Peers	Improve.	(all st.)	Peers	Improve.		
American Indian	4	4	4	1	5	4	1	4	4	3.44	0.33
Asian	6	4	5	4	5	4	5	4	6	4.78	0.56
Pacific Islander	4	4	4	1	5	4	1	4	4	3.44	0.22
Black	3	3	3	1	3	5	1	3	2	2.67	-1.00
Hispanic	3	4	4	1	4	4	1	4	4	3.22	-0.11
White	5	4	4	3	3	4	3	4	4	3.78	-0.22
Multi-racial	4	4	4	2	4	4	2	4	4	3.56	-0.22
Special education	2	4	4	1	4	4	1	2	1	2.56	-1.22
ELL	1	5	5	1	4	4	1	3	3	3.00	-0.11
Low income	4	5	5	1	5	4	1	4	4	3.67	0.22
Average	3.60	4.10	4.20	1.60	4.20	4.10	1.70	3.60	3.60	3.41	-0.16

NOTE: Ratings in red are less than the previous year, ratings in green are more than the previous year.

Using this modified matrix has a number of benefits. It preserves the simplicity of the Accountability Index matrix, uses the same metrics as the provisional Index to provide greater simplicity,⁸ provides more detailed information about subgroups, focuses on improvement from each group's baseline, relies on multiple cells when computing row average to reduce fluctuations in averages from year to year, and treats every group with equal weight regardless of the size of the group.

- Create an alternate method for the ELL group to make AYP by linking the results of the Washington Language Proficiency Test (WLPT) to the content assessments. *Schools and districts do not make AYP if the percentage of ELLs in WLPT Levels 2 and 3 who are on track to meet standard in the content areas (i.e., reading and math) when they become proficient in English declines two years in a row.*⁹ OSPI has developed a method to calculate the percentage of ELLs who are on track to meet the content standard when they become proficient in English.¹⁰ This alternative method is a fairer way to hold this group accountable because it emphasizes improvement in both English proficiency and academic performance and considers English language ability when examining students' performance in the academic subjects.

⁷ The current AYP system requires the use of unexcused absence rates at the elementary and middle school levels. Data for these rates are not included because they are not part of the index system, and nearly all schools meet the required goals.

⁸ The modified matrix relies on the same rules as the Accountability Index. For example, the results for all grades are combined, there must be at least 10 students to report results, there is no margin of error, the percent meeting standard is used for achievement indicator, the Learning Index is used for the peers and improvement indicators, and the same rating system is used.

⁹ OSPI has developed a method to calculate this percentage. This alternative method is a better way to hold this group accountable because it emphasizes improvement in both English proficiency and academic performance.

¹⁰ ELLs should be counted in WLPT Level 1 for only one year to provide an incentive to help new ELLs as much as possible.

Hence, the system is “compensatory” in nature—having one low rating in a matrix does not automatically result in a school/district not making AYP. The index blends performance across multiple ratings, and low ratings are compensated by higher ratings, a concept similar to how a GPA works. This is different from the “conjunctive” model now used to determine AYP. In a conjunctive model, a single missed target results in a school/district not making AYP. This is analogous to labeling a student as a failure when a single low grade occurs. The increasing level of proficiency currently required to make AYP will make it even less likely a school/district will meet the target. So in this analogy, a student would have to get higher and higher grades in all subjects to avoid being labeled a failure.

Consequences

State law currently prohibits the use of some consequences authorized by NCLB. For example, the state has no authority to require schools/districts to accept state help, and it has no authority to require corrective action or restructuring (e.g., remove staff, change curriculum, change governance). NCLB currently requires schools and districts to undergo increasing levels of “sanctions” if they do not make AYP over an extended period of time. NCLB also requires schools and districts that are in an improvement step to make AYP two years in a row in order to exit improvement status.

Many stakeholders believe these sanctions have flaws that need to be corrected. For example, students must be allowed to transfer to another school before their current school is required to provide additional services to help these students. Schools must also allow all students to transfer, including those performing well. Even when a school in “improvement” status makes AYP, all the sanctions must remain in effect. Finally, the consequences do not apply to non-Title I schools that do not make AYP, even though in Washington these schools outnumber Title I schools, enroll more students, and are more likely to have low index results (see Table 6). In fact, almost half the students in the state attended non-Title I schools that did not make AYP in 2008, and a large number of these schools are in “improvement” status but evade the teeth of the accountability system because they are not required to face any of the NCLB sanctions.

Table 6: Index Results for Schools Not Making AYP in 2008

Tier	Title I Schools		Non-Title I Schools	
	Number	Students enrolled	Number	Students enrolled
Exemplary	5	2,445	8	4,973
Very Good	27	13,211	31	30,637
Good	165	82,518	217	184,184
Fair	326	157,312	333	227,112
Struggling	56	19,184	83	24,388
Total	579	274,670	672	471,294

As a result, Washington proposes using a different set of consequences that reflect common sense changes to the current NCLB rules. These consequences should apply to all schools and districts, not just those receiving federal Title I funds.¹¹

¹¹ This should occur as long as the state does not lose any Title I funding due to federal “supplant” rules.

- Schools/districts not making AYP for the same reason (e.g., same subgroup) in consecutive years move into “improvement” unless there is a compelling reason not to, based on the results of a deeper review (see below).
- If the reason for not making AYP is due to the performance of a different group than a group responsible for not making AYP in the previous year, the school/district does not move to the next step of the process.
- School choice and supplemental educational services must be made available to the students in the subgroup(s) whose results are responsible for the school not making AYP. (Currently it applies to all students in the school, even if they are in a group that has performed well.)
- Those in improvement status that make AYP move back a step (e.g., from Step 2 to Step 1). Those making AYP two years in a row exit improvement status. This allows a gradual withdrawal of state support over time. (Under the current AYP rules, schools and districts in “improvement” must make AYP in two consecutive years to exit this status entirely, and no credit is given for making AYP in one year.)

Identifying Schools and Districts Needing Improvement

Each fall OSPI will compute the accountability index and apply the rules for making AYP. All schools and districts in all tiers will be given an AYP status, not just those receiving Title I funds. The first time a school or district does not make AYP, it is in a “warning” year. Schools and districts that do not make AYP two years in a row should not automatically fall into “improvement” status. Instead, they should undergo an in-depth review. The results of this review would determine if the school/district should move into an “improvement” step and be required to take certain actions.

The data to be reviewed fall in five general categories. The list below provides examples in each.

- **Contextual Data**
 - Type of school
 - Changes in student population
 - Programs served by the school
- **Assessment Results (State content assessments/WAAS/WLPT)**
 - Trends over multiple years for each subject area
 - Subgroup trends (e.g., race/ethnicity, ELL, special education)
 - Results for students who have been enrolled for at least two years
- **Teaching and Learning Issues**
 - Teacher education and experience levels
 - Student/teacher ratio
 - Recent changes in leadership (key central office staff and principals) and teachers
 - Alignment of curriculum and materials across grades and with state standards
- **Other Data**
 - Graduation and dropout rates for subgroups
 - Funding from local levies/bonds and outside sources
 - Problems with data that generate the index (e.g., reporting errors related to graduates)
 - Participation rates for all subgroups
 - Other indicator data (unexcused absence and graduation rates) for all subgroups
 - Perception survey results

Data will be closely reviewed for schools and districts that have not made AYP four years in a row, or meet other federal or state criteria. The state may determine that a school/district would benefit from a significant amount of additional support and move it to Voluntary Action for at least two years. If extra assistance is not accepted and significant improvement does not occur during the two-year period, the school would move to Required Action and a binding corrective action plan should be established between the district and the state, if authorized by the Legislature.¹² SBE and OSPI are working together to develop a process to identify schools that would move into Voluntary and Required Action in a way that conforms to newly emerging federal regulations. Moreover, the details of what assistance would be provided are still being developed.

When the details of the proposed system are finalized in the coming months, SBE and OSPI will submit a unified accountability plan to the U.S. Education Department that will recommend using the state accountability system for federal accountability purposes. A new administration may provide more flexibility to states that design alternative systems that provide more rigorous and valid accountability. All the data current reporting requirements of NCLB would continue to be met (i.e., making public the disaggregated data for the assessments, participation, and “other indicators” for the various student subgroups). Moreover, new data elements would be made public to further increase the rigor of the system.

Advantages Over the Current System

The proposed accountability system has many desirable features that make it a better alternative to the current rules used to measure AYP. The proposed state accountability system increases the system’s validity and rigor, reduces volatility and unintended consequences, makes the system easier to understand, supports the continued use of high standards and expectations, and provides more appropriate consequences when performance falls short of expectations.

- The Index is a *more valid* measure of school and district performance because it is based on the performance of all students in more subjects, is more nuanced than a Yes/No (pass/fail) system, and addresses several unintended consequences created by the current AYP system (e.g., narrowing the curriculum, focusing on students performing close to meeting standard).
- The Index is *more inclusive/comprehensive* because it uses a smaller minimum number for reporting (10 students across the entire school/district), includes the results of all students (not just those continuously enrolled through the testing period), includes both writing and science (this helps prevent a narrowed curriculum), and uses the Learning Index to measure performance across the range of assessment results (reduces the focus on “bubble” students who perform close to the proficiency cut point at the expense of students who are farther above and below that level).
- The Index is *less volatile* over time because assessment results are combined across all grades in a school and district (which raises the N) rather than using results for individual grades where students change from one year to the next.
- The Index is *more transparent/easier to understand* because it does not include a margin of error, the benchmarks are the same over time and among the different subjects for both schools and districts, there are fewer rules and “cells” to examine, and schools and districts have the same minimum number required for reporting the results.
- Using the Index to determine AYP helps the state *maintain high performance standards*. Two recent studies found that Washington has some of the nation’s toughest AYP requirements,

¹² ESHB 2261, passed by the 2009 Legislature, contains language on this issue.

resulting in a high percentage of schools not making AYP.¹³ Using the index in a compensatory manner reduces the incentive for the state to lower its standards so all students can be counted as proficient by 2014, a target viewed as unrealistic if standards are kept high.

- The proposed system has *more appropriate consequences* and provides *stronger intervention* measures when improvement does not occur.

RECOGNITION

Index results can be calculated retroactively and used for recognition purposes. Providing recognition based on 2009 results would be considered “Phase I” in the implementation of the accountability system, with full implementation contingent upon the provision of adequate funding. Using the index in this way will provide a more valid picture of school/district performance than AYP results, and it will introduce the concepts to the various stakeholders prior to its full implementation.

The recognition system should (1) be transparent and simple to understand, (2) rely mainly on criterion-referenced measures, and (3) provide multiple ways to demonstrate success. The recognition system is based on the belief that people are motivated more by success than by blame or guilt, and they need clear, challenging, and attainable goals.

SBE and OSPI are working together to create a unified recognition system based on the index results. SBE should give recognition for “Outstanding Overall Performance” while allowing OSPI to develop forms of recognition of their own. For example, OSPI could recognize a certain percentage of schools in math and science. Monetary compensation is not recommended, although matrix data could be used to generate schoolwide bonuses if the Legislature includes these as part of any law or reforms of the basic education finance system in the future.

Outstanding Overall Performance (7 types)

SBE has approved using the results from the Accountability Index to provide recognition when performance is very high. To ensure only truly outstanding performance is recognized, schools and districts must meet certain conditions. Theoretically all schools should be able to achieve recognition because it is a criterion-referenced system. Seven areas will be recognized for “Outstanding Overall Performance” using the following criteria.

(1) For the **index**, the 2-year average was at least 5.50, at least 10 cells of the matrix were rated each year, and there were fewer than 10% students designated as gifted each year.

(2-5) For **language arts** (reading and writing combined), **math**, **science**, and the **extended graduation rate**, the overall (column) 2-year average was at least 6.00, at least 2 of the 4 cells in the column were rated each year, and there were fewer than 10% students designated as gifted each year.¹⁴

(6) For the **achievement gap**, there were at least 10 students in at least 2 of the 5 outcomes (columns) in both of the income-related cells (non-low income and low income), there was no rating of 1 in any income-related cell or peer cell, there was no more than a 1-point difference in the rating

¹³ See “*The Accountability Illusion*,” Thomas B. Fordham Institute (February 2009) and “*Schools Struggling to Meet Key Goal on Accountability*,” Education Week (January 7, 2009).

¹⁴ For language arts, both reading and writing must have a 2-year average of at least 6.00 and at least 2 of the cells rated each year.

between the two income-related cells,¹⁵ and there were fewer than 10% students designated as gifted each year. Each of the above criteria had to be met two years in a row.

Each of the above six recognition areas requires fewer than 10% of the students to be gifted each year. Statewide approximately 3% of all students received this designation in 2008, so schools with 10% or more gifted students have unusually high concentrations of the most capable students. The exclusion criterion prevents a school from receiving recognition because it will likely have much higher than normal ratings. To ensure these types of schools are eligible to receive recognition, a seventh recognition area was established.

(7) For **schools with gifted programs**, the top 5% of schools in grade band—elementary, middle, high, and multi-level—that have at least 10% gifted students should receive this type of recognition, based on the 2-year average peer ratings in all four subjects.¹⁶

Table 7 shows the number of schools that would have been recognized by SBE in 2008. In all, 99 schools would have been recognized in 2008 if these criteria were in effect at that time, which is nearly 5% of all schools statewide. Of the 99, there were 8 alternative schools represented among the four school types.

Table 7: Distribution of Schools Meeting Recognition Criteria, by Grade Type (2008)
(Number of alternative schools in parentheses)

	Elementary	Middle/ Jr. High	High	Multiple Levels	Total*
Total Recognized	52 (1)	8 (3)	23 (0)	16 (4)	99 (8)
Number of All Schools	1,059	359	400	298	2,116
Percentage of All Schools Recognized	4.9%	2.2%	5.8%	5.4%	4.7%

CONCLUSION

President Obama has cited his concerns about the current AYP system, such as its failure to acknowledge when significant improvement has taken place. The President believes we can improve and evolve in all aspects of education, including how AYP decisions are made. We join him in his desire to change NCLB and the current AYP system in order to hold our schools and districts accountable in a more rigorous, more valid, and more transparent way. Washington has taken the initiative to lay out a new accountability model using a new index. Use of the index for making AYP decisions addresses fundamental weaknesses in the existing system and encourages the state to maintain rigorous content and performance standards. Stakeholders in Washington believe this new system and the use of the new index paves a way forward to increased clarity and accuracy in assessing our education system, thereby offering educators and stakeholders a transparent means to ensure each and every student receives an excellent and equitable education.

¹⁵ For example, if the reading non-low income cell is rated 5, the reading low-income cell could be rated no lower than 4 and no higher than 6.

¹⁶ Results for the peer indicators control for the types of students attending the school (the percent gifted, low income, ELL, special education, and mobile). This ensures schools with the highest concentrations of gifted students do not automatically receive this form of recognition.